

CLAIMS

We claim:

1. A computer-implemented method for performing an operation upon one or more graphs, wherein each graph represents a set of relationships between a set of biological molecules, wherein each graph comprises vertices representing the biological molecules and edges representing the relationships between the biological molecules, the method comprising
performing one or more operations on the one or more graphs to produce one or more product graphs.
2. The method of claim 1 wherein the operations comprise finding a common subset of vertices and edges in a plurality of graphs.
3. The method of claim 1 wherein the operations comprise merging a plurality of graphs having one or more common vertices or edges.
4. The method of claim 1 wherein the operations comprise deleting vertices and edges present in a first graph that are not present in a second graph.
5. The method of claim 1 wherein the operations comprise combining the edges and vertices of a plurality of graphs.
6. The method of claim 1 wherein the operations comprise finding a common subset of vertices and edges present in a predetermined percent of a plurality of graphs.
7. The method of claim 1 wherein the operations comprise finding a common subset of vertices and edges in a plurality of graphs, deleting the common subset of vertices and edges from each of the graphs to produce a plurality of graphs each with a unique set of vertices and edges.
8. The method of claim 1 wherein the operation is a recursive operation.
9. The method of claim 1 wherein the set of biological molecules comprises more than one type of biological molecule.
10. The method of claim 1 wherein the set of relationships comprises more than one type of relationship.
11. The method of claim 1 wherein at least one edge comprises an edge weight.

12. The method of claim 11 wherein the edge weight represents a value characterizing the relationship represented by the edge.
13. The method of claim 12 wherein the value is a numerical value.
14. The method of claim 11 wherein at least one edge comprises an edge weight table comprising the edge weight.
15. The method of claim 14 wherein the edge weight table further comprises one or more additional edge weights.
16. The method of claim 11 wherein at least one edge weight comprises an indication of a state.
17. The method of claim 11 wherein at least one edge weight comprises a spatial distance.
18. The method of claim 17 wherein the spatial distance represents a physical distance between the biological molecules represented by the vertices connected by the edge.
19. The method of claim 11 wherein at least one edge weight comprises a kinetic measurement.
20. The method of claim 11 wherein at least one edge weight comprises a distance metric representing a logical relationship between the biological molecules represented by the vertices connected by the edge.
21. The method of claim 11 wherein at least one edge weight comprises a statistical metric representing a logical relationship between the biological molecules represented by the vertices connected by the edge.
22. The method of claim 11 wherein at least one edge weight comprises a value of fuzzy set membership representing a logical relationship between the biological molecules represented by the vertices connected by the edge.
23. The method of claim 11 wherein at least one edge weight comprises a conditional probability.
24. The method of claim 23 wherein the conditional probability is the probability of a causal relationship between the biological molecules represented by the vertices connected by the edge.
25. The method of claim 1 wherein at least one edge comprises a direction.

26. The method of claim 1 wherein at least one edge comprises a boolean value indicating the presence or absence of an association between the biological molecules represented by the vertices connected by the edge.

27. The method of claim 26 wherein the association is co-expression, co-regulation, or presence or use in the same pathway.

28. The method of claim 1 wherein the biological molecules are selected from the group consisting of genes, open reading frames, expressed sequence tags, single nucleotide polymorphisms, sequence tag sites, nucleic acids, DNA, RNA, mRNA, cDNA, proteins, peptides, enzymes, metabolites, carbohydrates, exons, introns, cleavage fragments, restriction fragments, amino acid modifications, protein domains, DNA or RNA secondary or tertiary structures, nucleic acid motifs, protein motifs, and metal ions.

29. The method of claim 1 wherein at least two of the vertices represent different types of biological molecules.

30. The method of claim 1 wherein at least two edges represent different types of relationships between the biological molecules represented by the vertices connected by the edges.

31. The method of claim 1 wherein at least one edge represents a plurality of different types of relationships between the biological molecules represented by the vertices connected by the edge.

32. The method of claim 1 wherein the relationships are selected from the group consisting of physical distances between genes, open reading frames, single nucleotide polymorphisms, expressed sequence tags, sequence tag sites, or a combination thereof; genetic distances between genes, open reading frames, single nucleotide polymorphisms, expressed sequence tags, sequence tag sites, or a combination thereof; protein-protein interactions; protein-nucleic acid interactions; gene expression regulation; protein expression regulation; cellular signal transduction pathways; sequence similarity between genes or proteins; structural similarity between proteins; radiation hybrid mapping distances between genes, open reading frames, single nucleotide polymorphisms, expressed sequence tags, sequence tag sites, or a combination thereof; and metabolic pathways.

33. The method of claim 1 wherein at least one of the graphs comprises at least one hyper-edge.

34. The method of claim 33 wherein at least one of the operations converts at least one hyper-edge to a non-hyper-edge.

35. The method of claim 1 wherein at least one of the graphs comprises at least one hyper-vertex.

36. The method of claim 35 wherein at least one of the operations converts at least one hyper-vertex to a non-hyper-vertex.

37. The method of claim 1 wherein at least one of the graphs comprises at least one hyper-edge and at least one hyper-vertex.

38. The method of claim 37 wherein at least one of the operations converts at least one hyper-edge to a non-hyper-edge.

39. The method of claim 37 wherein at least one of the operations converts at least one hyper-vertex to a non-hyper-vertex.

40. The method of claim 37 wherein at least one of the operations converts at least one hyper-edge to a non-hyper-edge and at least one hyper-vertex to a non-hyper-vertex.

41. The method of claim 1 wherein at least one of the operations converts at least one edge to a hyper-edge.

42. The method of claim 41 wherein the hyper-edge is formed by combining two or more edges.

43. The method of claim 1 wherein at least one of the operations converts at least one vertex to a hyper-vertex.

44. The method of claim 43 wherein the hyper-vertex is formed by combining two or more vertices.

45. The method of claim 1 wherein at least one of the operations converts at least one edge to a hyper-edge and at least one vertex to a hyper-vertex.

46. The method of claim 45 wherein the hyper-edge is formed by combining two or more edges and the hyper-vertex is formed by combining two or more vertices.

47. The method of claim 1 wherein the product graph is modified relative to the graph on which the operation is performed.

48. The method of claim 1 wherein the operations comprise

delete all edges beyond a selected range of edge weights.

49. The method of claim 1 wherein the operations comprise dividing one graph into two graphs.

50. A computer-implemented method for performing an operation upon a graph, the graph representing relationships between biological molecules and having vertices representing the molecules and edges representing the relationships, the method comprising

- identifying a subset of zero or more of the edges,
- identifying a subset of zero or more of the vertices, and
- performing a unary operation upon the identified subset of edges and vertices to produce a product graph.

51. The method of claim 50 wherein the subset of edges identified are all edges beyond a selected range of edge weights.

52. A computer-implemented method for representing relationships between biological molecules using one or more graphs each having vertices and edges, the method comprising

- representing a set of biological molecules, wherein each molecule is represented by a vertex of the graph, and
- representing a set of relationships between the biological molecules, wherein each relationship is represented by an edge of the graph, wherein the edge connects two vertices,

wherein the graph is produced by performing one or more operations on one or more input graphs to produce the one or more graphs.

53. A computer program product for performing an operation upon one or more graphs, wherein each graph represents a set of relationships between a set of biological molecules, wherein each graph comprises vertices representing the biological molecules and edges representing the relationships between the biological molecules, the computer program product comprising a computer data medium on which is carried

- a means for performing one or more operations on the one or more graphs to produce one or more product graphs.

54. A computer program product for performing an operation upon a graph, the graph representing relationships between biological molecules and having vertices

representing the molecules and edges representing the relationships, the computer program product comprising a computer data medium on which is carried

a means for identifying a subset of zero or more of the edges,

a means for identifying a subset of zero or more of the vertices, and

a means for performing a unary operation upon the identified subset of edges and vertices to produce a product graph.

55. A computer program product for representing relationships between biological molecules using a graph having vertices and edges, the computer program product comprising a computer data medium on which is carried

a means for representing a set of biological molecules, wherein each molecule is represented by a vertex of the graph, and

a means for representing a set of relationships between the biological molecules, wherein each relationship is represented by an edge of the graph, wherein the edge connects two vertices.

56. A computer-implemented method for representing relationships between biological molecules using a graph having vertices and edges, the method comprising

representing a set of biological molecules, wherein each molecule is represented by a vertex of the graph, and

representing a set of relationships between the biological molecules, wherein each relationship is represented by an edge of the graph, wherein the edge connects two vertices.

57. A representation of relationships between biological molecules comprising one or more graphs each having vertices and edges, each graph comprising

a set of biological molecules, wherein each molecule is represented by a vertex of the graph, and

a set of relationships between the biological molecules, wherein each relationship is represented by an edge of the graph, wherein the edge connects two vertices,

wherein the graph is produced by performing one or more operations on one or more input graphs to produce the one or more graphs.

58. The representation of claim 57 wherein the set of biological molecules comprises more than one type of biological molecule.

59. The representation of claim 57 wherein the set of relationships comprises more than one type of relationship.

60. A data structure comprising a representation of relationships between biological molecules, the representation comprising a graph having vertices and edges, the graph comprising

a set of biological molecules, wherein each molecule is represented by a vertex of the graph, and

a set of relationships between the biological molecules, wherein each relationship is represented by an edge of the graph, wherein the edge connects two vertices.

61. A computer-implemented method for performing an operation upon one or more graphs, wherein each graph represents a set of relationships between a set of biological molecules, wherein each graph comprises vertices representing the biological molecules and edges representing the relationships between the biological molecules, wherein the biological molecules, the relationships between the biological molecules, or both, are derived from different sources, the method comprising

performing one or more operations on the one or more graphs to produce one or more product graphs.

62. A computer-implemented method for performing an operation upon one or more graphs, wherein each graph represents a set of relationships between a set of biological molecules, wherein each graph comprises vertices representing the biological molecules and edges representing the relationships between the biological molecules,

wherein at least two of the vertices represent different types of biological molecules, at least two edges represent different types of relationships between the biological molecules represented by the vertices connected by the edges, at least one edge represents a plurality of different types of relationships between the biological molecules represented by the vertices connected by the edge, at least one vertex represents a plurality of different types of biological molecules, or a combination thereof,

the method comprising

performing one or more operations on the one or more graphs to produce one or more product graphs.

63. A computer-implemented method for performing an operation upon one or more graphs, wherein each graph represents a set of relationships between a set of biological molecules, wherein each graph comprises vertices representing the biological molecules and edges representing the relationships between the biological molecules, wherein the biological molecules, the relationships between the biological molecules, or both, are derived from heterogeneous molecular biological data, the method comprising

performing one or more operations on the one or more graphs to produce one or more product graphs.